United States Department of Agriculture

Service Center Modernization Initiative (SCMI)

Standard for Geospatial Dataset File Naming

Prepared by
Data Management Team #5: Geospatial Data Standards

Abstract: This document provides the USDA Service Center Modernization Initiative standard for geospatial directory and file naming conventions. It describes the conventions used for the basic nationally consistent set of core geospatial data, locally acquired geospatial data and derived geospatial data.

Keywords: geospatial data, GIS, file naming, standard

Published and distributed on January 25, 2001
(Revision of Standard for Geospatial Dataset File Naming: SCI Std 004-01, August 31, 2000)

Document number: SCMI Std 004-02

Introduction

As directed by the Secretary of Agriculture's March 16, 1998 memorandum, the Natural Resources Conservation Service (NRCS), Farm Service Agency (FSA), and Rural Development (RD) agencies are co-locating offices, modernizing business processes, and partnering to achieve a "one-stop service" for United States Department of Agriculture (USDA) customers at their county-based field offices (Service Centers). One of the major components of the modernization initiative involves the implementation of a Geographic Information System (GIS) across each of the Partner Agencies and in all 2,550 Service Center offices. A Service Center Data Team has been chartered with the overall responsibility for implementing an infrastructure for management of data resources for the Partner Agencies. The GIS Standards Team 5 was formed to address specific data management issues regarding geospatial data.

The individuals who contributed to the development of this standard are:

David Anderson, (NRCS) Service Center Data Team Leader
Carol Ernst, (FSA) Co-Leader
Emil Horvath, (NRCS) Co-Leader

Liz Cook (NRCS)	Rodney Johnson (FSA)	Ron Selph (NRCS)
Dwain Daniels (NRCS)	Dave Nabidy (FSA)	Phil Teague (NRCS)
Larry Davis (NRCS)	Steve Nechero (NRCS)	Rob Vreeland (NRCS)
Randy English (NRCS)	Elaine Ortiz (NRCS)	Nicole Soltyka(SAIC)
Kent Williams (FSA)	Jill Schuler (NRCS)	Randy Frosh (Unisys)

Working group list

ARC/INFO is a registered trademark of Environmental Systems Research Institute, Inc.
ArcView is a registered trademark of Environmental Systems Research Institute, Inc.
Microsoft Access 97 is a registered trademark of Microsoft Corporation.
Microsoft Access 2000 is a registered trademark of Microsoft Corporation.
MrSid is a registered trademark of LizardTech, Inc.

Table of contents

This page is intentionally left blank

STANDARD FOR GEOSPATIAL DATASET FILE NAMING

1. Overview

The objectives of this standard are to help in managing United States Department of Agriculture (USDA) Service Center Modernization Initiative (SCMI) geospatial data by establishing directory (folder) and file naming conventions; support the concurrent USDA Service Center Modernization Strategy to develop a basic nationally consistent set of core geospatial data that will provide a foundation on which to base business applications; and to relate to other SCMI geospatial standards including SCMI Std 003, *Standard for Geospatial Data Set Metadata* [A2][1], SCMI Std 005, *Standard for Geospatial Feature Metadata* [A3], SCMI Std 007, *Standard for Geospatial Data* [A4], and the *USDA Service Center Initiative Directory Structure and File Naming Convention Change Control Policy* [A6].

This standard contains two appendices. Appendix A provides bibliography references. Appendix B contains the initial directory and naming conventions established for the Business Process Reengineering (BPR) pilot sites.

1.1. Scope

The scope of this standard is to define the directory and file naming conventions for the *geospatial dataset collection* (physical repository of data) that resides at a Service Center. This standard shall apply to the set of nationally consistent core geospatial data layers first defined in the *USDA Service Center Geographic Information System (GIS) Strategy* [A5]. It also provides standards on the directory structure and file naming for locally acquired and derived geospatial data. This document replaces the initial directory and naming conventions established for BPR pilot sites known as version 5 (See Appendix B).

1.2. Purpose

GIS for the Service Center is expected to comprise nationwide coverage of more than 20 common *geospatial datasets* (a group of similar spatial phenomena) that are collected and distributed at the county level of geography. To organize this data at the Service Center so that it is accessible, maintainable and updateable requires a standard scheme for categorizing the data into directories and establishing names and conventions for the files in the directories.

This standard will continue to evolve as nationally consistent datasets are provided to the Service Centers. However, this document is an initial attempt to identify the directories and file names for existing common geospatial dataset categories and it helps to establish initial standards and direction for BPR projects that require geospatial data.

[1] The number in brackets corresponds to those of the bibliography in Appendix A.

This standard will be placed under configuration management and maintained through a structured change control process because the impact of changing this standard can be great on those applications that use the data and those who provide the data. The change control process will allow proposed changes to be reviewed and discussed by those affected by the change.

Nationally fielded applications will be developed that rely on the nationally consistent set of geospatial data. These applications will rely on the integrity of the data in meeting the specifications in this standard. Applications that are built locally for a Service Center or for data that is acquired locally shall adhere to these standards.

1.3. Acronyms and abbreviations

BPR	Business Process Reengineering
CCE	Common Computing Environment
CD-ROM	Compact Disc Read Only Memory
CLU	Common Land Unit
DEM	Digital Elevation Model
DLU	District Land Unit
DMF	Digital Map Finishing
DOQ	Digital Ortho Quadrangle
DOQQ	Digital Ortho Quarter Quadrangle
DRG	Digital Raster Graphs
EPA	Environmental Protection Agency
FEMA	Federal Emergency Management Agency
FIPS	Federal Information Processing Standard
FSA	Farm Service Agency
FWS	Fish and Wildlife Service
GIS	Geographic Information System
GNIS	Geographic Names Information System
ISO	International Standards Organization
MDOQ	Mosaicked Digital Ortho Quadrangles
MLRA	Major Land Resource Areas
MrSID	Multi-resolution Seamless Image Database
NAPP	National Aerial Photography Program
NASIS	National Soil Information System
NCGC	National Cartography and Geospatial Center
NRCS	Natural Resources Conservation Service
NWI	National Wetland Inventory
OIP	Office Information Profile
RD	Rural Development
SSA	Soil Survey Area
SCMI	Service Center Modernization Initiative
SSURGO	Soil Survey Geographic Database
STSSAID	State Soil Survey Area ID
TIF	Tagged Image File
US	United States

USDA	United States Department of Agriculture
USGS	United States Geological Survey
UTM	Universal Transverse Mercator
WRP	Wetland Reserve Program

2. Background

The *USDA Service Center Geographic Information System (GIS) Strategy* [A5] first defined a list of geospatial datasets required to provide a foundation on which to base business applications. The *Geospatial Data Acquisition, Integration, and Delivery National Implementation Strategy Plan* [A1] further refined and expanded this list. This list was organized into logical categories based on business names. It is these logical categories that form the basis for the organization of the physical directory structure defined in this standard.

The SCMI Std 007, *Standard for Geospatial Data* [A4] includes a geospatial data model that details a hierarchical classification that shall be used to categorize, or provide taxonomy for, geospatial data. These categories are referred to as *geospatial dataset categories.* This model and classification shall be used to identify and describe geospatial data in a consistent way. The *geospatial dataset categories* in the model are used as the basis for the directory structure in this document.

The top level of the geospatial data model is referred to as a *geospatial dataset collection.* A *geospatial dataset collection* is a catalog and physical repository of *geospatial datasets*. For example, an USDA data mart that serves geospatial data to only one Service Center, or an USDA national data warehouse that serves geospatial data to all Service Centers is a *geospatial dataset collection.*

This standard is based on the lessons learned during the initial fielding of geospatial datasets at the pilot sites (see Appendix B). In addition to creating a flatter directory structure, each dataset name within each geospatial dataset category was examined in order to provide a consistent naming convention that would offer a standard method of dataset identification including name, data type and location. The directory structure and naming conventions resulting from this examination follow in the subsequent sections. The geospatial data model is used as the basis for the directory structure in this document.

3. Geospatial dataset collection

The entire *geospatial dataset collection* at the top level of the directory shall be located on a designated drive and named "geodata". This replaces the previous top-level directory "Service Center Themes", the sub-directory named according to Service Center Office Information Profile (OIP) number and name, and the sub-directories within that named according to county (see Appendix B). Removing the Service Center and county level sub-directory level simplified navigation but has ramifications on geospatial dataset file naming which are discussed in 5.1. There shall be only be one “geodata” occurrence in

any given Service Center. The standard directory structure has been reduced and flattened as follows:

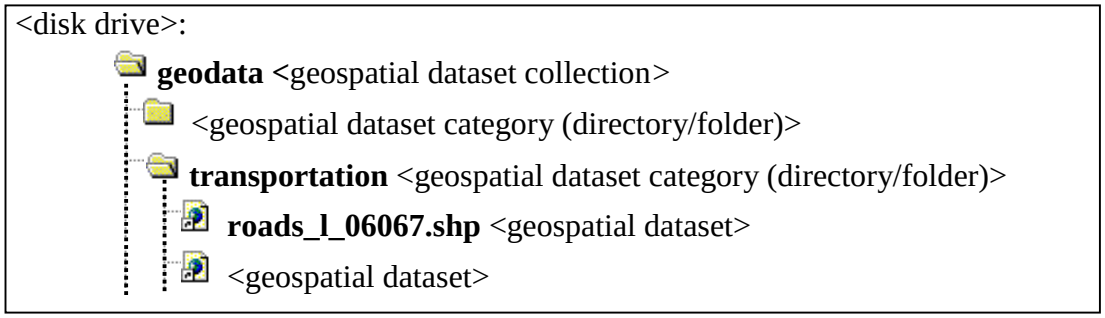


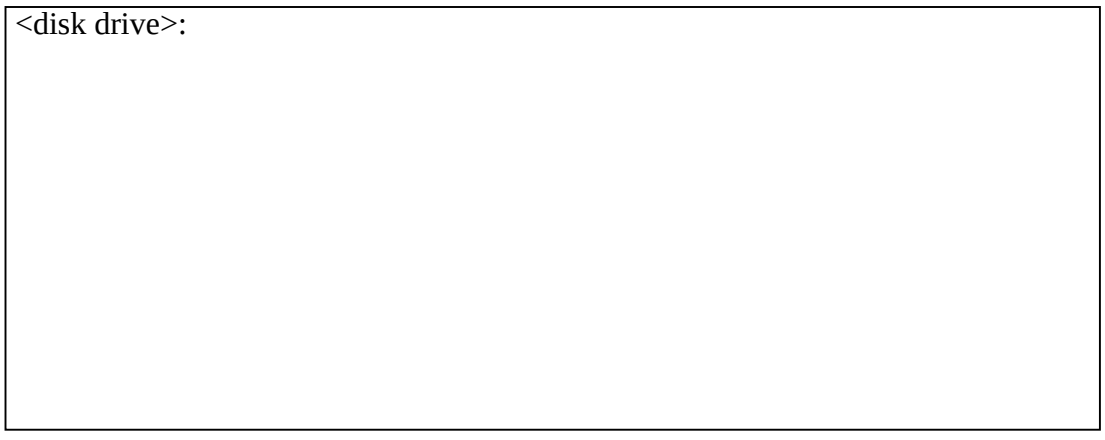
Figure 1—Geospatial dataset collection

An example of the physical path to a road map using this standard on NT would be
C:\geodata\transportation\roads_l_06067.shp
On the Service Center server the physical path would be F:\…

For UNIX it would be
/geodata/transportation/roads_l_06067.shp

4. Geospatial dataset category

The next level in the geospatial data model is a *geospatial dataset category, which* is a logical group or division of a *geospatial dataset collection.* A *geospatial dataset category* is analogous to a computer directory or folder. There can be multiple occurrences of *geospatial dataset* categories under the *geospatial dataset* collection ("geodata") level. A *geospatial dataset category* may include at least one *geospatial dataset* complete with all metadata and feature data including geometry, attributes, labels and symbology. However, an empty directory can exist as a 'place holder' for future data. Additionally, *geospatial dataset categories* are allowed to have sub-directories as in the case of climate.



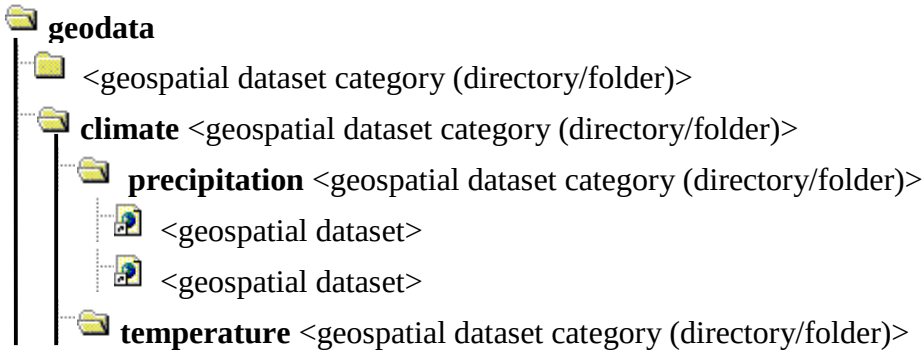


Figure 2— Geospatial dataset categories

4.1. Geospatial dataset category naming

In this second level data hierarchy, the physical directory name previously used upper case characters and spaces. In practical applications these two practices cause problems migrating datasets between UNIX and NT platforms. As a result, this standard requires that the physical implementation of the category directory names shall consist of **only** lower case characters "a-z". Additionally, the underscore character "_" shall be used in place of a space in a directory name. Allowable characters are covered in detail in Clause 6.

The SCMI Std 007, *Standard for Geospatial Data* [A4] identifies the *geospatial dataset categories* used in this standard. These standard category or directory names are also shown in Table 1. The standard name used for each directory hopefully shall reflect a name that is commonly used and understood by Service Center personnel when referring to the *geospatial dataset category*. For example, *geospatial dataset categories* include common_land_unit, ortho_imagery and soils (See Table 1).

5. Geospatial dataset

A *geospatial dataset* is a group of similar spatial phenomena in a *geospatial dataset category* and is related to one metadata set. A *geospatial dataset* is often referred to as a layer, theme, coverage, or simply a map. For example, the *geospatial dataset category* hydrography could contain the *geospatial datasets* surface water, water control infrastructure and flood hazard maps. The surface water *geospatial dataset* contains streams represented as lines, ponds represented as polygons and wells represented as points.

5.1. Geospatial dataset naming

Table 1 shows the standard file name for each dataset. Table 1 also shows a dataset title that users might commonly use to refer to the dataset and a description of the dataset. The names are designed to be unique within the entire geodata directory. They maintain their uniqueness even if the category or directory names are eliminated from the structure. The

standard file names convey as much information as possible and reflect encoding into the name of

- dataset theme
- type of map features in the dataset
- spatial location or extent of the dataset

As a result, the dataset name contains

- a short version or acronym used to represent the business name of the dataset
- feature type designation p-point, l-line, a-area, t-text, g-grid, r-raster, e-enhanced Digital Raster Graphs (DRG), s-Multi-resolution Seamless Image Database (MrSID), d-database/excel, i-index
- location or extent information such as a Federal Information Processing Standard (FIPS) code or State Soil Survey Area ID (STSSAID) number

6. Standard characters

Because producers and consumers of geospatial data use computers with different operating systems, (e.g., NT, UNIX) several restrictions must be imposed on naming files to ensure all systems can access the data. The application software often places other restrictions.

6.1. Special characters

In an effort to design a list of allowable characters both computer platform and GIS software filename restrictions must be considered. In terms of operating systems, NT does not allow the characters "\ / : * ? < >" in file names. Use of a "." can be problematic in NT as it indicates a file suffix.

UNIX allows any character but some of the characters in the preceding NT list can be problematic. In addition using spaces in names or beginning a name with "-" is problematic for UNIX. Beginning a name with "." in UNIX indicates a hidden file and requires additional switches to the command that produces a directory listing.

The GIS currently in operation at the BPR sites allows dashes "-" and underscores "_" but not periods ".". The full function GIS requires that a coverage name begins with a letter.

6.2. Case sensitivity

Use of upper and lower case characters in names is common and very useful in UNIX. However, this can be problematic with NT because it does not distinguish between a file named for instance "FileName" and "filename". On UNIX, the full function GIS converts all names to lowercase. The desktop GIS converts all characters to lower case for a new shapefile name.

6.3. Allowable standard characters

The allowable standard characters identified in this standard are based on the least common denominator for both operating systems and software. This approach will not impede any potential migration of *geospatial datasets* to a new platform or new software environment as technology and software enhancements are realized in the future.

Because of the combination of all these restrictions the **only** characters allowed in a standard file name are the following:

- lower case a-z
- the numerals 0-9
- the underscore "_" character
- the dash "-" charter

In addition, the first character shall be a letter a-z.

These restrictions also shall apply to *geospatial category* or directory naming. These standards are very restrictive and shall not change unless the GIS platform changes with further definition from the Common Computing Environment (CCE) Team.

7. Name length

The total length of the dataset filename shall not exceed 30 characters. This limit has been identified during Compact Disc Read Only Memory (CD-ROM) production at National Cartography and Geospatial Center (NCGC). The only CD-ROM writing format that is universally readable throughout Natural Resources Conservation Service (NRCS) is the International Standards Organization (ISO) 9660, Level 2, Mode 1 format. Exceeding this 30-character maximum for filenames becomes a problem when datasets are sent via CD-ROM to multiple computer platforms.

However, no attempt is made to adhere to the so-called 8.3 format required by older DOS operating systems (maximum of eight character name with a maximum of a three character suffix). This will cause problems for Microsoft Access 97 and Access 2000 because they can not import or link to .dbf files (such as those in a shapefile set) whose names are longer than 8.3 until they are renamed with an 8.3 compliant name. (Note: Access requires 8.3 for imported and exported dbf files even though it will handle the longer names.)

There is a 13 character maximum for coverages in the ARC/INFO GIS. This was exceeded in this standard since it is designed to organize shapefiles and would prove to be easier to use in the Service Centers because it will lessen the need for cryptic names unfamiliar to many Service Center users. However, conversion of data either to or from coverages will require different names and additional processing.

8. Area of interest

As mentioned previously the spatial location or extent is encoded into the name. In the current scheme most maps will be clipped or tiled to the county boundary for delivery and use unless the map is a state or national coverage. This clipped extent shall be appended to the theme name and feature type as a FIPS code or STSSAID number (e.g., roads_l_**06048**.shp).

Tiling of digital geospatial data significantly impacts overall data management and system performance. In general, it is preferred that tiling is seamless, or transparent, to the user.

Clipping map extents at the county boundary is not optimal for users if they for instance wish to look at a farm or an area that crosses a county boundary. Users would prefer to be able to zoom to some arbitrary area of interest and remain unaware of the underlying database structure or tiling scheme. However, this is not possible given the current state of technology.

As a result, clipping the map data to the county boundary is the best available option for delivery and maintenance of map data. This standard's encoded file naming scheme and directory structure ensures that maps of like datasets in adjacent counties will appear next to each other in a pick list to facilitate user selection of maps.

In some cases, such as Digital Raster Graphs (DRG) and Digital Ortho Quadrangles (DOQ), an image catalog serves as an index map to 7.5-minute quad tiles for a county. This index is used to display images and conceals the underlying tiling scheme and image filenames from the user.

9. Local data

Any GIS data that is acquired or developed locally at the Service Center shall be placed in the geodata directory along with its completed metadata. A few guidelines are offered to assist in the incorporation of this data in a logical and consistent manner.

9.1. Existing "geospatial dataset category"

If the data corresponds to one of the existing categories or directories, the map and metadata should be placed in that directory so that it appears adjacent to any nationally provided data when the user is selecting from a pick list.

There are several circumstances where various types of local Service Center data should be incorporated into the existing "geospatial dataset category" structure. They are

- **When there is no national data and only locally developed data**, such as in the case of Wetland Reserve Program (WRP) easements data. These maps along with the Metadata should be placed in the environmental_easements directory or other relevant "geospatial dataset category" directories.

— **When there are more accurate data from local sources** the locally acquired data should be placed in the appropriate "geospatial dataset category" directories. For instance, there is locally obtained road data that are known to be more accurate than the nationally provided data named roads_l_<stnnn>. The locally acquired data and its metadata should be placed in the transportation directory and uniquely named according to the standard. One approach is to encode the datasets scale into the filename. For instance, roads12k_l_<stnnn>, indicates that the road map is locally acquired 1:12,000 scale data.

— **When data are created locally as the result of analysis** it should be placed under the appropriate "geospatial dataset category" directory if the data would be useful to others at a Service Center. The data should be named according to the standard and placed in the appropriate geodata directory. Otherwise, the results should be left on a personal disk drive where it is probably inaccessible to others.

9.2. "Local_geodata" catchall directory

When the *geospatial dataset category* of the locally acquired data does not fit in any of the existing categories, the data should be placed in the catchall directory named "local_geodata" in the "geodata" directory.

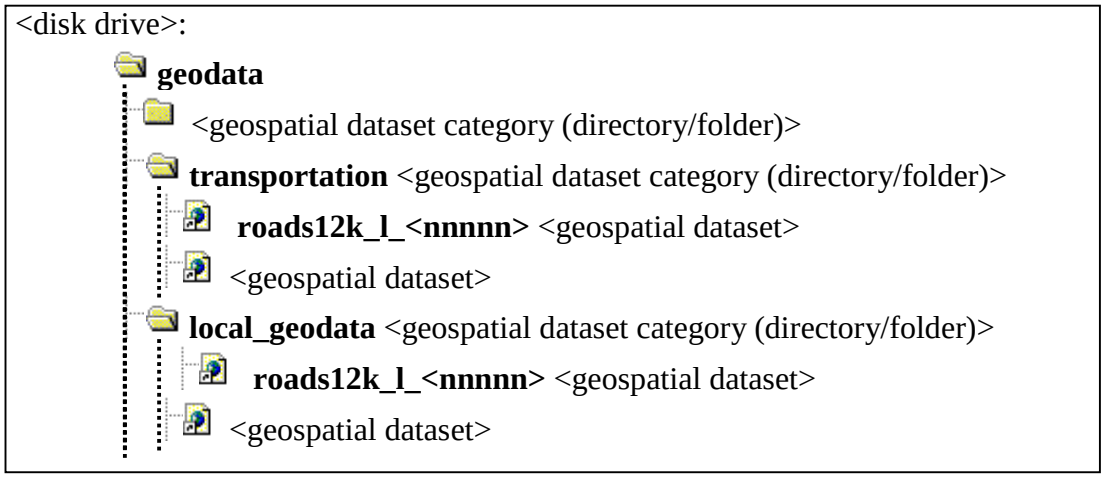


Figure 3— Geospatial data structure with sub-directory "local_geodata"

9.3. Local data naming

Any locally acquired data should be named in accordance with this standard and only lower case a-z, 0-9 and "_" characters used in names as described in 6.3. Also, the theme, feature type and spatial location or extent should be encoded into the name as described in 5.1. Additionally, the name must not exceed 30 characters as described in Clause 7.

10. Geospatial data directory structure: "geodata"

This section and table describes map layers common to all BPR sites and Service Centers.

10.1. Overview

The physical implementation of directories and file names supports the USDA Service Center Modernization Strategy to develop a basic nationally consistent set of core geospatial data that will provide a foundation on which to base business applications.

10.2. Table notations

Table 1 identifies the specific categories and geospatial files within each category and the standard naming conventions for the file. The table applies the naming guidelines set forth in 5.1 of this standard where each file name encodes: the dataset theme, feature type, and location for which the dataset applies.

The following notations apply to the file naming conventions used in Table 1:

- < > indicates a substitution notation
- **( )** indicates a choice list notation
- | indicates a choice of options and reads as "or"
- **<mmm>** is the three-letter abbreviation for the applicable month (e.g., precip<mmm>_a_<st>, precipjun_a_co is the file name for Colorado June precipitation)
- **<nnn>** is the congress number
- **<stnnn>** is the 2-character state postal abbreviation and 3-digit County FIPS codes (e.g., drg_r_<mdnnn>, drg_s_md047 is the file name for Worcester County, Maryland DRG)
- **<a>** is a substitution for the leading character that describes the Mosaicked Digital Ortho Quadrangles (MDOQ) imagery
- m indicates DOQQs (Digitial Ortho Quarter Quadrangles) are present and reside in native Universal Transverse Mercator (UTM) zone
- x indicates there is a missing DOQQ in the DOQ
- z represents re-projected Digital Ortho Quadrangle (DOQ) into dominant county UTM zone
- **<nnnnnnn>** is a 2-digit latitude number, 3-digit longitude number, and 2-digit quadrangle number (e.g., <a><nnnnnnn>, m3010601 is a Tagged Image Format (TIF) DOQ for the native UTM zone and 30 north latitude 106 west longitude, sheet 1 of 64)
- **<st>** is the two character state postal abbreviation (e.g., precip_a_<st>; precip_a_co is the filename for Colorado annual precipitation)
- **<stssaid>** is the state soil survey area ID (e.g., soils_l_<stssaid>; soils_l_24047 is the filename for Worcester County, Maryland Soil Survey Geographic Database (SSURGO) Lines)
- **<usgs standard>** is the standard naming convention used by the United States Geological Survey (USGS)
- **<x-x>** is number - total tiles in a county ortho mosaic. Tiles are numbered west to east and north to south. These are county subsets due to maximum file sizes;
- **<none>** is no file name yet assigned because data has never yet been delivered
- **<County Name>** is the actual name of the county for the dataset title

- **us** indicates a dataset covering United States
- **p, l, a,** or **t** indicate the dataset feature types of point, line, area, or table in a shapefile set (.shp .dbf .shx .sbn .sbx .prj)
- **g** indicates a grid dataset feature type
- **r** indicates a raster dataset feature type
- **s** indicates a MrSID compression raster dataset feature type
- **e** indicates enhanced DRG with map collar removed
- **i** indicates an image catalog dataset feature type (.dbf);
- **d** indicates a database/excel spreadsheet dataset feature type

Tab1e 1—Geospatial data "geodata" directory structure and naming conventions

Directory	File name	Dataset title	Description
air_quality	<none>	Air Quality <County Name>	No files delivered to date
cadastral	ntlparks_a_<stnnn>	National Park Areas <County Name>	National Park Polygon data
	ntlparks_l_<stnnn>	National Park Boundaries <County Name>	National Park Line boundaries for cartographic display
	plss_a_<stnnn>	Public Land Survey System <County Name>	Public Land Survey System polygon data
	plss_l_<stnnn>	Township Range Section <County Name>	Public Land Survey System boundaries for cartographic display. No files delivered to date
census			(Demographics) Defined Later – Identify options to procure data
climate\precipitation	precip_a_<st>	Annual Precipitation Range < State>	'precipitation' is a subdirectory of 'climate'. Annual precipitation (sum of 12 monthly maps) for the entire state. <st> is equal to the state two character postal abbreviation
	precip_l_<st>	Annual Precipitation Isolines <State>	Annual precipitation boundaries for cartographic display for the entire state. No files delivered to date
	precip<mmm>_a_<st>	<mmm> Precipitation Range <State>	Mean (1961-1990) Monthly precipitation data for the entire state. <mmm> is equal to the three-letter abbreviation for the applicable month
	precip<mmm>_l_<st>	<mmm> Precipitation Isolines <State>	No files delivered to date
climate\temperature	<none>		'temperature ' is a subdirectory of 'climate'. No files delivered for Service Center to date.
common_land_unit	clu_a_<stnnn>	CLU <County Name>	Common Land Unit (CLU) – Farm field boundary
	dlu_a_<stnnn>	DLU <County Name>	District Land Unit (DLU) – Farm field boundary

Directory	File name	Dataset title	Description
conservation_practices	<none>	Planned and Applied Conservation Practices	Conservation practice data aggregated for the Service Center. Toolkit group will resolve naming. No files delivered to date. Data is developed locally
cultural_resources	<none>	Cultural Resources <County Name>	Archeology, state historic sites, Native American settlements and burial grounds, National Park Service National Register of Historic Places, National Historic Landmarks and National Natural Landmarks. No files delivered to date
elevation	contour_l_<stnnn>	Contour Lines <County Name>	1:24,000 USGS hypsography line data
	ngs_p_<stnnn>	Geodetic Survey Monuments <County Name>	Location and description of National Geodetic Survey Monuments (point data)
	ned_r_<st>	NED <State Name>	1:24,000 USGS National Elevation Dataset (NED) merged into a seamless raster format with elevations portrayed in meters.
	<usgs standard - native format>	DEM<Full Quad Name>	1:24,000 USGS Digital Elevation Model (DEM). USGS standard lat/long name with a “d” leading character
endangered_habitat	<none>		No files delivered to date.
environmental_easements	wrp_a_st	Wetland Reserve Program <State Name>	Aggregation of WRP easements for State Service Centers. No files delivered to date. Data developed locally
	wrp_a_stnnn	Wetland Reserve Program <County Name>	Aggregation of WRP easements for a specific Service Center area. No files delivered to date. Data developed locally
geographic_names	gnis_p_<stnnn>	Topo Place Names <County Name>	Geographic Names Information Systems point data from GNIS concise file
govenment_units	boundary_l_<stnnn>	Administrative Boundaries <County Name>	1:24,000 USGS boundary line data

Directory	File name	Dataset title	Description
	boundary_a_<stnnn>	Administrative Areas <County Name>	1:24,000 USGS boundary polygon data (state park, wildlife refuge, etc.)
	congdist_<nnn>_a_st	Congressional District <nnn>	Full US Congressional districts 104 – 106. <nnn> is the congress number
	cities_p_<stnnn>	Cities <County Name>	Cities point data from Geographic Names Information System (GNIS) populated places file
	cnty24K_a_<stnnn>	County Boundary <County Name>	1:24,000 county boundary polygon data
	cnty24K _l_<stnnn>	County Line <County Name>	1:24,000 county boundaries for cartographic display
.	cnty100k_a_<stnnn>	County Boundary <County Name>	1:100,000 county boundary polygon data from Census Tiger data.
	cnty100k_l_<stnnn>	County Line <County Name>	1:100,000 county boundary for cartographic display from Census Tiger data.
	manfetr_a_<stnnn>	Man Made Area Features <County Name>	1:24,000 USGS manmade feature polygon
	manfetr_l_<stnnn>	Man Made Line Features <County Name>	1:24,000 USGS manmade feature line data
	rcd_a_us	Resource Conservation & Development Areas	Full US Resource Conservation & Development Areas polygon data
	state_a_us	State Areas	Full US state polygons
	state_l_us	State Boundaries	Full US state boundaries for cartographic display
	swcd _a_st	Soil and Water Conservation District	Full US Soil and Water Conservation District polygon data
	zip_p_us	Zip Codes	Full US zip code centroids (points). GIS Implementation Team to identify data source
hydrography	damsites_p_<stnnn>	National Inventory of Dams <County Name>	National Inventory of Dams point data
	femaq3_a_<stnnn>	Flood Hazard Maps (FEMA) <County Name>	Federal Emergency Management Agency (FEMA) polygon data
	hydro_dmf_l_<stnnn>	Rivers and Streams (DMF) <County Name>	1:24,000 Soil Survey Digital Map Finishing (DMF) line data
	hydro_l_<stnnn>	Rivers and Streams <County Name>	1:24,000 USGS line data

Directory	File name	Dataset title	Description
	hydro_rf_l_<stnnn>	Rivers and Streams (EPA Reach 3) <County Name>	1:100,000 Environmental Protection Agency (EPA) Reach File line data
	ntlhydro_l_<stnnn>	Rivers and Streams <County Name>	1:100,000 USGS/EPA National Hydrography Dataset line data
hydrologic_units	huc24K_a_<stnnn>	Hydrologic Units <County Name>	1:24,000 polygon data of the Hydrologic Units at the 5th and 6th level.
imagery	<none>		Other imagery files such as satellite or non-standard imagery. No files delivered to date
land_use_land_cover	lulc_a_<stnnn>	Land Use Land Cover <County Name>	Polygon data of the Land Use Land Cover
	nonveg_a_<stnnn>	Barren Land (Topo Map) <County Name>	1:24,000 USGS non-vegetative polygon data (sand area, beach, gravel beach, etc.)
	nlcd_r_<st>	Land Use Land Cover <State Name>	30 meter USGS/EPA National Land Cover Dataset raster data.
	surfcvr_a_<stnnn>	Surface Cover (Topo Map) <County Name>	1:24,000 USGS surface cover polygon data (woods, brush, orchard, etc.)
map_indexes	napp_p_<stnnn>	Photo Index (NAPP) <County Name>	National Aerial Photography Program (NAPP) point data
	quads12k_a_<stnnn>	Quarterquad Areas <County Name>	1:12,000 quad polygon data
	quads12k_l_<stnnn>	Quarterquad Lines <County Name>	1:12,000 quad boundaries for cartographic display. No files delivered to date
	quads24k_a_<stnnn>	Quadrangle Areas <County Name>	1:24,000 quad polygons
	quads24k_l_<stnnn>	Quadrangle Lines <County Name>	1:24,000 quad boundaries for cartographic display. No files delivered to date
ortho_imagery	mosaic<x-x>_s_<stnnn>	County ortho <County Name>	APFO MrSID county ortho mosaic of MDOQ. <x-x> number-total titles.
	ortho_i_<stnnn>	Ortho Image Catalog <County Name>	Ortho image catalog in DbaseIV format for TIF DOQs

Directory	File name	Dataset title	Description
	<a><nnnnnnn>	MDOQ <Full Quad Name>	TIF DOQ <a><nnnnnnn> is leading character, two spaces for latitude, three spaces for longitude and two spaces for the 01 to 64 quadrangle numbers in the one degree block. Leading character can equal: m – all DOQQs present and reside in native UTM zone x – there is a missing DOQQ in the DOQ z – re-projected DOQ into dominant county UTM zone
	<usgs standard>	DOQQ <Quarterquad Name>	USGS Quarterquads – Any format (.bil, .bsq, .bip, .tif). Text name of quarter quads with "_" for imbedded spaces.
	mosaic_s_<stnnn>	County ortho <County Name>	NCGC MrSID county ortho mosaic
plants	<none>		No files delivered to date
soils	crpdata_d_<stssaid>	1990 CRP Frozen Soil List	NOT A MAP-Excel spreadsheet with 1990 frozen soils data used for Conservation Reserve Program (CRP) eligibility determinations
	soilattributes_d_<stssaid>	<Soil Survey Area Name> Soil Attributes	NOT A MAP-Access data base of National Soil Information System (NASIS) distribution format 1.0
	mlra_a_us	Major Land Resource Areas	Full US Polygon data of Major Land Resource Areas (MLRA) Reselected to SC Area
	soil_a_<stssaid>	SSURGO Areas <Survey Area Name>	SSURGO Soils Polygon data
	soil_l_<stssaid>	SSURGO Lines <Survey Area Name>	Outlines of the SSURGO polygon boundaries for cartographic display
	soil_p_<stssaid>	SSURGO Points <Survey Area Name>	Point data of the soils special features
	soillfetr_l_<stssaid>	SSURGO Linear Features <Survey Area Name>	Line data of the soils special features

Directory	File name	Dataset title	Description
	ssa_a_<stssaid>	SSA Boundary <Soil Survey Name>	Polygon data limit of Soil Survey Area (SSA)
topographic_images	drg_i_<stnnn>	Topo Image Catalog <County Name>	TIF Digital Raster Graph index of enhanced DRG
	drg_s_<stnnn>	Topo <County Name>	MrSID Digital Raster Graphs without map collar
	<usgs standard>	<Full Quad Name> - Topo	USGS DRG Quad with collar
	<usgs standard>_e	<Full Quad Name> - Enhanced Topo	Enhanced DRG image with map collar removed
transportation	misctrans_l_<stnnn>	Utility lines (Topo Map) <County Name>	1:24,000 USGS line data (power transmission lines, substation, pipelines, etc.)
	railroads_l_<stnnn>	Railroads (Topo Map) <County Name>	1:24,000 USGS line data-railroad layer
	railroads_dmf_l_<stnnn>	Railroads (DMF) <County Name>	1:24,000 Soil Survey (Digital Map Finishing) railroad line data
	roads_l_<stnnn>	Roads (Topo Map) <County Name>	1:24,000 USGS line data-Roads layer
	roads_dmf_l_<stnnn>	Roads (DMF) <County Name>	1:24,000 Soil Survey (Digital Map Finishing) roads line data
wetlands	nwi_a_<stnnn>	NWI (FWS) <County Name>	Polygon data of the National Wetland Inventory (NWI) Fish and Wildlife Service (FWS)
	nwilfetr_l_<stnnn>	NWI Linear Features (FWS) <County Name>	Linear Features line data of the NWI
	nwi_l_<stnnn>	NWI Lines (FWS) <County Name>	Outlines of the NWI polygon data for cartographic display
	nwi_p_<stnnn>	NWI Points (FWS) <County Name>	Point data of the NW I

Appendix A – Bibliography

When the following standards are superseded by an approved revision, the revision shall apply.

[A1] Geospatial Data Acquisition, Integration, and Delivery National Implementation Strategy Plan, Draft #4 Service Center Business Process Reengineering Data AID Team, September 22, 1999

[A2] SCMI Std 003, Standard for Geospatial Data Set Metadata

[A3] SCMI Std 005, Standard for Geospatial Feature Metadata [This standard is currently under development.]

[A4] SCMI Std 007, Standard for Geospatial Data

[A5] USDA Service Center Geographic Information System (GIS) Strategy, Interagency Team, August 18, 1998

[A6] USDA Service Center Initiative Directory Structure and File Naming Convention Change Control Policy, Initial Draft, IO Lab, October 8, 1999

Appendix B – Business Process Reengineering directory structure

Prior to the development of this standard, the initial directory structures (referred to as version 5) for the *geospatial dataset collection* fielded at the Business Process Reengineering (BPR) sites reflected the storage structure at the data warehouse that supplied the data. In order to maintain all county based geospatial datasets in a single warehouse separated according to Service Center, the directory structure was defined as:

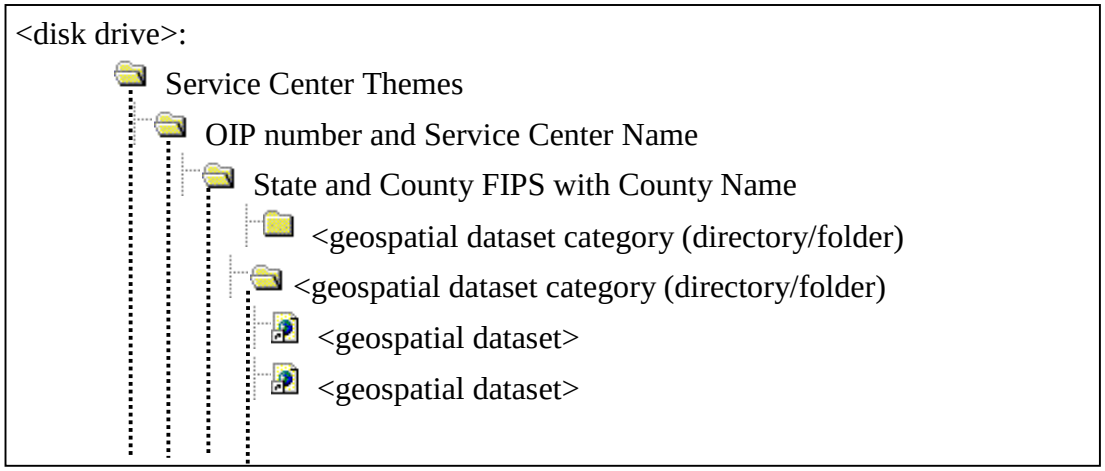


Figure B.1—Initial directory structure

This structure organizes datasets within a generic Service Center Themes folder according to a unique Service Center Office Information Profile (OIP) number and second according to county. Each county serviced within a Service Center possesses its own uniquely identified folder (directory) and is named according to the five-digit Federal Information Processing Standard (FIPS) code and county name. For instance, the actual path to a roads map on NT would be:

C:\Service Center Themes\2487 Sacramento\06067 Sacramento\Transportation\roads.shp

Personnel at the BPR sites using the desktop GIS, ArcView in its native mode to access geospatial datasets, found the length and depth of this directory structure too cumbersome for accessing geospatial datasets. For example, Service Centers that service only one county still had to navigate through the county level directory (<State and County FIPS with County Name>) to access dataset categories where only one county's worth of data existed. Similarly, the Service Center level of the directory structure (<OIP Number and Service Center Name>) had to be negotiated even though there would only ever be one occurrence at this level. As a result of this learning experience at the BPR pilot sites and experiences developing applications that have to traverse this directory structure, the directory structure was flattened and streamlined to facilitate access to the datasets with fewer steps.